PADCO/U.S.A.I.D. SOCIAL SECTOR RESTRUCTURING PROGRAM
PADCO International Inc., Frolovskaya 6/8 Corpus 1, Kyiv, Ukraine, 254070
Tel: 38 044 463 7615, 16, 17, 18, 19: FAX 38 044 463 7620
e-mail: rvaughan@padco.kiev.ua

CREATING A NATIONAL PENSIONER DATA BASE

AND

IMPROVING THE PENSION REPORTING SYSTEM

Prepared by PADCO Social Sector Reform Project

October, 1998

TABLE OF CONTENTS

INTRODUCTION

Ukraine should create a national pensioner database. This database should contain information on all Ukrainians who are currently receiving pension benefits – including information about their address, age, sex, benefits (current and paid in the past) and the factors used to determine pension benefits. The pension system and the evaluation of reform options require sophisticated models and accurate computations. And accurate computations requires accurate input data. Small errors in assumptions caused by lack of information can result in unnecessary errors in financial and actuarial projections.

The cost of developing and maintaining the pensioner database is small in relation to its value. This is especially true in Ukraine, since such databases have been created and maintained for the last 10-15 years or more at the raion level. Only minor efforts are required, therefore, to bring together raion data files, and reformat and join them into a nationwide database.

Reasons for creating a national pension database are:

- The pension reform needs accurate calculations. The Pension Database will be the unique source for accurate data about existing pensioners. Financial projections of the pension system will be based, at least in part, on this database
- The Pension Databases already exist at the raion level, and, in some parts of Ukraine, at the oblast level. There is also a national standard for raion pensioner databases. As part of our ongoing actuarial support to the government of Ukraine, we have analyzed databases from Mykolaiv and Volyn oblasts. From this, we have already identified many of the problems that must be solved to create a national database
- National Pensioner Database analysis can be used in two ways. The set of resulting tables can be used as a source for periodic Pension Statistical Bulletins as well as for input to the government's pension analysis models. It can also be used as a source for critical pension system management information.
- The National Pensioner Database is a critical step in the creation of an integrated system of national social assistance databases. This is important because many social programs are interrelated. For example, social pensions impact the calculation and eligibility for housing subsidies. The Ministry of Labor and Social Policy also needs to have information about individuals participating in more than one social program
- An evaluation of the required hardware to support the National Pensioner Database shows only limited special equipment will be needed. Even a modern ordinary PC can maintain it. However, a workstation with a 9-GB hard drive, 3 to 4 gigabytes of RAM, and a Unix operating system is needed to handle and analyze a database of this size.

WHY SHOULD UKRAINE IMPLEMENT THE NATIONAL PENSIONER DATABASE NOW?

Favorable conditions now exist in Ukraine for the creation and maintenance of a National Pensioner Database. Ukraine has begun the process of pension reform. The first legislative package has been drafted. As the Cabinet of Ministers proceeds from concepts to concrete designs and implementation, the need for accurate data to support financial projections will become greater.

Paragraph thirteen of the supplement to the Decree of the Cabinet of Ministers, dated 6.27.98, states: *"With the purpose of carrying out actuarial valuations, substantiating and estimating social-economic consequences of implemented decisions and draft laws to improve a statistical recording and reporting in the sphere of pension security, allowing for implementation of necessary data for estimating the current state of the pension system and expenses to cover pension payments."* The proposed National Pensioner database should be the one source of such information.

Consequently, on the one hand, policy makers want to make well-thought out and carefully calculated decisions. On the other hand, local departments of social protection may provide them with the necessary data. In fact, the Ministry of Labor and Social Policy, along with the oblast state administrations are joint owners of the pensioner database. It is characterized by:

- Electronic data storage. Pensioner data is stored in electronic form as database records;
- Completeness. This database covers ALL pensioners in Ukraine (apart from a small number of career military personnel not covered by social protection agencies);
- Universality. The database stores data on every pensioner and contains more than 30 individual characteristics (last name, address, sex, age, earnings, period of employment, pension benefit, social benefits, additional payments, bonuses, number of dependents, etc.).
- Inconsistent data storage. Currently, pension data is stored in raion databases in a variety of different formats. There isn't a single computer program to maintain all these databases, but there is a single data standard with everyone should comply with. Format incompatibility is a major technical obstacle that impedes the creation of the National Pensioner Database. Thus, enforcement of a unified standard is ones of the keys to its creation.

Today Ukraine possesses both the data required for accurate estimations and the desire to use them to improve financial forecasts. Only minor organizational and technical efforts are required to convert existing data into a format suitable for its implementation.

WHY DO WE NEED A NATIONAL PENSIONER DATABASE?

AT THE NATIONAL LEVEL

Creating a database at the national level will allow the GOU to:

- Improve pension legislation by allowing for more accurate information and financial forecasts;
- Improve statistical reporting;
- Perform special analyses required by policymakers and administrators;
- Coordinate social protection information systems;
- Carry out statistical and econometric surveys, classifying pensioners by various characteristics and combinations of characteristics[1];
- Simulate the impact of alternative policy decisions on different classes of pensioners and on the fiscal position of the pension fund;
- Calculate the impact of draft normative documents.

The last two points are critical, and require further discussion. For example, it is possible on a regular basis:

- To estimate the cost of:
  - Both social pensions and social components of standard pensions[2], and changes in how they are financed[3];
  - Privileged pensions (according to List No 1, etc.), and changes in how they are financed
  - Pension benefit increases (for example, to the to the minimum consumption level) for particular pension categories or for all pensioners.
- To evaluate alternative pension indexing decisions, and changes in the relationship between average wage and average pensions
- To provide accurate fiscal analysis of norms or legislative acts on supplements and benefits.
- To create population forecasts for the pension segment of the population. To use such forecasts for the simulation of various types of pension systems.
- By implementing economic and demographic forecasts it is possible:
  - To estimate consequences of introducing particular mechanisms for increasing pensions for various categories of pensioners;
  - To project changes in pension payments due to the changing composition of the work force, between industry and agriculture, and among industries.[5] For example, as the number of war veterans declines, supplements payable to this group will decline. Or if there is a reduction in the number of workers in the coal industry, this can impact the amount of privileged pensions payable. [Footnotes don't correspond!!]

[1] For example, by age, sex, category of an early retirement, etc.
[2] For example, an increase for caring for a person over 80 years, etc.
[3] For such calculations a supplementary information is required on organizational structures that financed pension payments in the past and will do it in the future.

- To estimate the costs of different mechanisms of pension recalculation or its indexing due to inflation, wage increases, and other parameters related to the functioning of the pension system.

- Introduction of the National Pensioner Database will improve management of the pension system in general, apart from improving data analysis. The quality of statistical reporting would be increased, because all reports would be generated in the National Pensioner Database center using automated, computerized procedures. Reporting will also become much more versatile. At present, these problems are resolved by primitive methods, if they are resolved at all.

AT THE OBLAST LEVEL

It is anticipated that the National Pensioner Database will be created by joining together the oblast pensioner databases. Thus, each oblast will maintain a "small" copy of the National Pensioner Database, with the same structure. In principle, oblast surveys are possible based on the same analysis modules used at the national level. If needed to satisfy the needs of the oblast level surveys, it's possible to develop supplementary analysis modules.

The Pensioner Database may become the backbone of an Integrated Social Protection Database. The Integrated database will be a source of information for an oblast department of information and analysis.

Creation of such a database may facilitate integration of different departmental oblast computer centers and their evolution into an independent oblast department of information and analysis. Such an approach can change the function of the computer center from mechanical data collection, to analysis and financial projections.

IN SUMMARY

- The National Pensioner Database can be used to answer the following major questions:
  - What pension benefits are being paid to selected groups of pensioners, or to specific individuals
  - How was the specific pension amount calculated? The database can give detailed information about wages, service history, privileged employment and supplements, for example
  - What changes in pensions are likely in the future? For example, the expected number of new retires between 2003 and 2007 will decrease due to the large number of deaths in the Great Patriotic War and funds will become available for other purposes, such as increasing the minimum or maximum pension
- Results of the analysis can be presented:
  - In the form of ready-made tables in a statistical report
  - As statistical reports generated in response to a specific query.
  - In the form of input data to for financial and actuarial models, or for econometric surveys.
- To analyze a draft law or a decision some extra effort will be necessary:
  - To include economic and demographic forecasts based on macroeconomic indicators. Note that, because of the Pensioner Database, analysis some indicators may be specified in regard to the pension system. ??

- To develop models. Results may be calculated based on data retrieved from the Database. A model may also be optimized by analyzing hidden data nuances that will be discovered in the process of database analysis.
- To create mechanisms to increase pension benefits and evaluate the cost of these measures.

These calculations can be highly accurate. The cost of creating and maintaining such a database will be insignificant compared to the cost of calculation and evaluation errors resulting from continuing to use the present decentralized sources of information.

HOW WILL THE NATIONAL DATABASE BE CREATED?

The National Pensioner Database will be a regularly updated data extract from raion pensioner databases. This means certain selected basic information about all pensioners will be extracted from raion pensioner databases to create the national database. The national database will contain some, but not all, information on all pensioners. All oblast departments of social protection will send such data extracts to the Ministry of Labor and Social Policy once a quarter. Raion databases will continue to be updated on a daily basis, and oblast databases will be updated on a monthly basis. The Ministry will simply combine all raion database extracts into the National Pensioner Database each quarter.[4]

Special software, a prototype of which was developed for analysis of the Mykolaiv and Volyn oblast pensioner databases, will be used for database inquiries and will generate results in tables ready for printing. Results could also be used as input data files, for example, for the financial and actuarial model of the national pension system.

PADCO experts can:

- Assist experts from the Ministry of Labor and Social Policy and the Pension Fund of Ukraine with the development of special software to analyze the National Pensioner Database and create standard statistical reports
- Work with experts from the Ministry of Labor and Social Policy, the Ministry of Finance, and the Pension Fund of Ukraine to develop financial and actuarial models that will use results of inquiries to the National Pensioner Database
- Consult with the GOU in the use and maintenance of the software programs, and transfer these functions to the GOU over time.

[4] Technically, it is advisable to use this retrieval to automate the preparation of quarterly oblast report No 94, when most of oblasts already accumulate raion databases.

WHAT ARE THE OBSTACLES TO CREATING A NATIONAL DATABASE?

A number of organizational and technical problems should be resolved to create the National Pensioner Database.

ORGANIZATIONAL TASKS

- Determine who will own the National Pensioner Database. Today, the Ministry of Labor and Social Policy owns raion databases. It is advisable that one of its subdivisions collect and process information. However, the Ministry of Finance, Ministry of Statistics, and Pension Fund should all have direct access to the database, and be able to generate queries and reports on their own
- Require that raion departments of social protection delete data concerning deceased pensioners from corresponding databases less frequently than they generate reports. In general, it is advisable to archive and keep such data at least for 30-50 years[5].
- Prepare normative documents to authorize the project. In general, this will be a package of instructions and covering letters from the Ministry of Labor and Social Policy concerning data extracts and their transfer, maintenance of pre-defined lists of terms, and ensuring validity, completeness and authenticity of data.
- Conclusion of an agreement of cooperation between PADCO and the Ministry on developing the National Pensioner Database and methods of access to existing raion databases.

There are several ways to implement the project. Two of them are.

- The "top-down" method. A database is created by order of administrative executives Different executives may have different interests. For example:
  - Deputy Minister Mikolya Noshchenko may need an estimate of the cost of transferring social pensions to the "NASHA SIMYA" caseload management computer software program.
  - Deputy Minister Olena Garyacha may use the database to estimate the overlapping needs of pensioners for different social assistance programs.
  - Deputy Minister Volodymyr Matvychuk may use the database as a way of analyzing the costs of alternative pension reform scenarios.
  - Chairman Boris Zaichuk might be most interested in using the model to assist theOffice of the Actuary with financial analysis of the solidarity system.
  - There are many other potential users (DerzhKomStat, etc.).
- The "bottom-up" method. An oblast administration may want to analyze data for their own needs. Mykolaiv, for example, needs the database to compile an oblast social budget.

PADCO may study such needs and develop supplementary analysis modules to use at the oblast level.

TECHNICAL TASKS

- Develop specifications for data stored in the National Pensioner Database.

[5] An accurate calculation of the expected lifetime of a particular person is critical for pension systems based on individual accounts. Duration of life of a particular person depends on many factors including heredity, quality of the health care system, environmental conditions, and economic status.

- Develop technical directions and documentation for retrieving data from an existing pensioner database maintained by raion departments of social protection.
- Develop standard output tables and reports based on the National Pensioner Database
- Develop a monthly or quarterly standard Statistical Report using automatically generated tables
- Develop an algorithm for database inquiries and corresponding software support.
- Develop a procedure for delivering the Oblast Pension Databese from the Oblast to the Center.
- Evaluate the best software platform to use for database analysis and reports. The could be based in part on a survey of software currently used by computer centers in oblast departments of social protection
- Define project implementation procedures, and evaluate potential oblasts for pilot implementation of the project.

WHO WILL MAINTAIN THE NATIONAL PENSIONER DATABASE? HOW MUCH WILL IT COST?

DATABASE POTENTIAL OWNER

The following organizations are potential database owners:

- Information and Computation Center of the Ministry of Labor and Social Policy;
- Monitoring and Research department of the office of complex analysis of the Ministry of Labor and Social Policy;
- Pension Reform Information and Analysis department recently created to maintain the National Pension Database under the Ministry of Labor and Social Policy.
- Pension Fund of Ukraine
- Office of the Actuary recently created under the Pension Fund of Ukraine.

Regardless of who is selected as the owner, other Ministries and organizations within the GOU must have easy access to the database, and be able to generate queries and reports on their own.

REQUIRED HARDWARE

The table shown below summarizes these requirements:

Project stages	**Minimum requirements (50% project controllability)**		**Requirements to an adequate functioning of the system (for a problem to be resolved in an adequate time period)**	
	Description	**Cost**	**Description**	**Cost**
Maintenance	Office computer Operating system Peripheral equipment	$1,500 $0.00 (freeware from Internet) $1,000	Workstation + 2 PC Operating system + DBMS Peripheral equipment	$15,000+$3,000 $3,000+3,000 $10,000
Program of forming an oblast data extract	Using PADCO's developed technical requirements and authorized by the decree of the Ministry of Labor, local programmers develop their own software for every oblast	$0.00 (developed by programmers of the Ministry of Labor and local programmers)	2 programmers working 1 week in each oblast (6 months for 25 oblasts of Ukraine)	2x6x$600= $7,200
Database	A form of the Pension Statistical report and modules are developed by PADCO staff			

analysis modules	

Figures mentioned above are presented for the purpose of evaluation and will be refined when the project is divided into stages. Each stage will be supported by its own implementation plan.

LOOKING AHEAD

The ultimate goal is to create a series of public databases based on proper operating systems and software, and using a standardized data format. The structure and composition of these data bases would be published and updated each time the database is changed so different government organizations are aware of all existing databases, their structure, and the information they contain. It should be possible for any public entity to submit query requests and quickly obtain information from any other database. For example, the Pension Fund should be able to get unemployment information from the unemployment center for use in pension calculations.

The component parts of such an information structure will include:

- A system of reporting, i.e. a system of periodic reporting of personified information regarding new retirees, deaths, and changes in pensioner status
- A system of recording, i.e. databases, that will be maintained (or are already maintained) by different public offices
- A system of information standards that will be included (or are already included) in the above mentioned databases.

A system of personified data reporting should allow for the vertical flow of information from a legal entity/individual to corresponding public offices. And a system of standards and public information should make separate department databases "transparent" to all users and facilitate horizontal data flows. This will allow different public agencies to synchronize and verify database information. This would be especially valuable to the State Tax Administration.

The National Pensioner Database and the personified database of all workers will be two of the most important components of the Integrated Data System. The knowledge gained from the design of these two systems can be generalized to the design of the whole system. It is true that at present neither raion nor oblast pensioner databases are unified. Thus, standardization of distributed raion/oblast databases is mandatory for the creation of the Integrated National Pensioner Database.

If the elements and structure of different databases are public information, then the data becomes more useful, and its value increases. Thus, a requirement for establishing links between the same data in different databases is an important condition for the creation of an integrated reporting system. This problem is resolved automatically for data collected and compiled through a standardized reporting system. Thus, information enters department databases from an information-processing center with already established links. The procedure is more complicated for data that was already in the database before development of the data reporting system. Links among raion databases are initially established, similar to a procedure that housing subsidy offices follow to use information stored in pension databases. These links along with the data are transferred from the bottom up to the oblast and national levels. The earlier mentioned system of standards will play a key role in the process of integration of existing databases.

It will take several steps to create the information structure.

- Defining standards and implementing common codes based on analysis of the structure of existing databases. These designs will be implemented both for establishing "horizontal" links among existing databases and creating an adequate system of collection and communication of data by data processing centers.

- Establishing links among databases within the same department.
- Establishing links among databases in different public offices.
- Creating a shell program for accessing integrated databases in on-line mode. Developing a data exchange protocol for integrated databases in the off-line mode.
- Monitoring and simulating based on data stored in integrated databases.

For all these reasons, the ultimate goal of shared and integrated databases should be kept in mind when designing and implementing the National Pensioner Database.

TECHNICAL APPENDIX

What is the minimum configuration of computers capable of efficiently operating with the National Database? The National Database will likely need to maintain data on every Ukrainian pensioner. Each record contains approximately thirty numeric parameters. Only 6 bytes are needed to store a number with 14-digit accuracy. Thus, data on a single pensioner will take about 200 bytes. In this case, information covering 15 million pensioners occupies about 3 gigabytes. It's possible to purchase ordinary computers with hard drives of this size.

Access to such a database using a PC does not depend on the speed of the processor, but will be determined by the speed of hard disk access speed. $200 hard disks currently ensure a read-out speed of about 2-Mb per second with a mean access time of 16 milliseconds. Consequently, 3 gigabytes could theoretically be read sequentially in about 25 minutes.

owever, most queries and reports will require random access to 15 million records, and will require about 66 hours to process. Only in rare circumstances, where the records can be accessed in the same order in which they exist in the database, will it take less than half an hour to process the request. All other requests could take close to three days. This might be acceptable for certain types of standard reports where calculation speed is not critical. However, such a speed is not adequate for time critical work characterized by a great number of complex inquiries. Please note that this proposal does not provide for the use of enhanced data storage techniques, nor does it anticipate on-line access to data. These features would substantially increase the overall cost of the system. If the GOU ultimately decides to include confidential information in the data base or use the data for calculations, system enhancements will be required.

The low cost of this database creation is that the figures it requires already exist and they need only to be collected and reformatted. Also, its low maintenance cost is explained by the following:

- Information will be input in the database in off-line mode on a quarterly basis. This is much cheaper than to maintain the database in on-line mode. We do not need neither costly telecommunication equipment, nor connections to high-speed communication networks, nor expensive hardware and software to support the database under such conditions. Several inexpensive modems and access to the database through the Internet or a similar local network will make it work at an acceptable level.
- The database will contain only information which is not confidential, so data security is less of an issue.. Therefore, we do not need to purchase expensive hard and soft ware to secure the data and authorize access for their users.
- The data can be saved using a simple packing function. And there is no need to purchase expensive input/output systems which provide parallel copying, expansion capacity, "hot" exchange, or automatic data recovery should any fault occur.
- The database can be based on a widely-used software and hardware platform. Thus, we do not need to employ "expensive" specialists to maintain exotic computer devices.

For these reasons, it's possible to configure a computer to handle data bases of this size within a reasonable time period at modest cost. Such a database may be analyzed by a modern desktop high-level

computer[6]. A specialized workstation costing between $8,000 and $15,000 would comfortably handle this database. Such a computer will have a special operating system (most likely Unix) which costs about $3,000. Alternatively, Linux, a variation of the same operating system may be found on the Internet as freeware.

In addition, the computer would have 3-4 gigabytes of Random Access Memory (RAM). The hard disk capacity must be sufficient to update the database. We are planning to do these updates once a quarter as oblast offices send quarterly data retrievals. The database volume will grow proportionally, due to new retirees, changes in retirement status, or benefit recalculations. Consequently the number of new or changed records will be 1,000,000 to 3,000,000 per year . Consequently, the annual increase in the database will be about 1/2 gigabyte. Even so, we can still use a standard disk sub-system instead of a more effective but expensive system such as RAID 1-5, etc. And some packing accessories costing approximately $1,000 will enable us to solve the problem of information saving.

Modems installed by PADCO in oblast departments of social protection are sufficient for data communication. Each oblast will send 2-6 megabytes per quarter to the Center. In the future, it's possible to allow on-line access to the database. For this purpose, we have to define a topological model of the system, purchase network and communication equipment, provide high-speed WAN connections, maintain software modernization, etc. However, this approach will make the project ten times more expensive, so such option should not be considered at this time.

Corresponding software will need be also developed. Since some software specifications were developed and refined in the process of creating the Mykolaiv and Volyn oblast pension databases, this software should require approximately 3-6 person months to refine.

The task of developing software for data retrieval can be solved in several different ways, two of which are descried below.

- Based on PADCO's technical requirements, the local designers of the current pension calculation programs develop corresponding software. In order to control the process, a corresponding decree of the Ministry should back it up.
- A mobile group of 3-4 qualified programmers visit each oblast, and with the approval of the Ministry of Labor and Social Policy, examine local data formats and design corresponding software support. PADCO already employs some of these top professionals.

[6] With an account of constantly reducing costs of hard disks and their growing sizes and also growing PC capacity, these computers will turn within six months into a lower level hardware.